California Postsecondary Education Commission

An Analytic Method for Measuring Student Academic Preparation

June 2008
FS 08-04

Background

One of the goals in the Commission's Performance Assessment Framework for higher education is that all high school students in California receive an education that gives them the opportunity to enter a college or university. There are many possible measures that could be used to gauge preparation for higher education. These include SAT scores, number of Advanced Placement (AP) courses taken, standardized test scores, and completion rates of the *a–g* college preparatory curriculum.

It is not always clear from the data which of these measures is most closely related to progress in higher education. Commission staff are developing an analytic method that looks at the relationship between the different measures available and progress in higher education. This analysis will help staff in selecting a measure that is the best gauge of movement toward the goal of giving students from all backgrounds the opportunity for higher education.

This fact sheet reports progress on a study exploring academic proficiency scores of public high school students, as these measures are linked directly to the knowledge and skills that students need to acquire to be successful at the university level. Staff found that a measure based on scores in high school proficiency tests correlates slightly better with university student persistence than does the school's Academic Performance Index (API). A distinct advantage of the Commission's measure is that it can be calculated by gender and by ethnic/racial group within schools. Thus, it can be used as an indicator of how well schools are doing in giving all students an education of a quality that prepares them for higher education. API is available only for a school as a whole, so it is not a good indicator of differences in the quality of preparation within a school.

The Commission's Academic Preparation Measure

A numerical scale that provides a general indication of the overall proficiency level of public high school students in selected college-preparatory subjects. The scale, which ranges from 0 to 900, correlates with one-year and two-year university persistence rates.

Staff Analysis

A key research focus was to develop a measure of academic preparation that would show a relationship with various measures of postsecondary learning and achievement. Commission staff derived subject-specific proficiency scores for each public high school by dividing the number of students who tested at or above proficiency in a subject by the number of test takers. Data were obtained from the California Standardized Test (CST) files that are maintained by the California Department of Education.

Staff correlated the CST academic proficiency model with two important measures of student success: one-year and two-year persistence rates for freshmen entering California State University and University of California. The one-year persistence rate is the proportion of an entering freshman class that re-enrolled for the subsequent fall term. The two-year persistence rate is the proportion still enrolled two years later.

Proficiency in one area is likely to be correlated with proficiency in other areas. Correlations between subject areas ranged from moderately high to very high, as shown in Display 1. For example, proficiency scores in Algebra I are highly correlated with proficiency scores in Geometry, Integrated Math 2, Algebra II, and World History. Proficiency in Algebra I is also correlated with proficiency in Chemistry, Earth Science, and Physics, although the correlation is not as high.

DISPLAY 1 Correlation Matrix of Academic Proficiency Scores

	Algebra I	Geometry	Integrated Math 2	Algebra II	World History	Biology/ Life Sci.	Chemistry	Earth Science	Physics
Algebra I	NA	0.88148	0.75241	0.73096	0.72434	0.70157	0.57568	0.58261	0.52163
Geometry	0.88148	NA	0.11886	0.82808	0.77271	0.77395	0.74214	0.58984	0.65483
Integrated Math 2	0.75241	0.11886	NA	0.20039	0.5825	0.57875	0.55158	0.29756	0.13607
Algebra II	0.73096	0.82808	0.20039	NA	0.7167	0.68936	0.71919	0.48968	0.64451
World History	0.72434	0.77271	0.5825	0.7167	NA	0.87342	0.72591	0.67184	0.64442
Biology/Life Sci.	0.70157	0.77395	0.57875	0.68936	0.87342	NA	0.73914	0.68875	0.63751
Chemistry	0.57568	0.74214	0.55158	0.71919	0.72591	0.73914	NA	0.54688	0.62561
Earth Science	0.58261	0.58984	0.29756	0.48968	0.67184	0.68875	0.54688	NA	0.53248
Physics	0.52163	0.65483	0.13607	0.64451	0.64442	0.63751	0.62561	0.53248	NA

A value close to -1 indicates a high negative correlation and a value close to +1 indicates a high positive correlation.

The statistical procedure derived weights for scores in each subject, based on the correlation matrix. The resulting single function represents student academic preparation as a combination of individual scores, as shown in Display 2.

To assess the validity of the model, staff correlated it with the API developed by the California Department of Education. Because both measures rely on academic proficiency, they are likely to be highly correlated. A low correlation value would raise serious validity issues. API scores and the Commission's CST proficiency score model correlated at an exceptionally high value of 0.88.

The preliminary analyses presented in Display 3 confirm that API and the Commission's linear model of academic proficiency are both associated with freshman persistence rates for CSU and UC, but the Commission's measure had slightly higher correlations and has the advantage that it can be calculated by gender and by ethnic/racial groups within schools. All correlations are statistically different from zero. Staff will be able to refine and modify this measure over time to increase the degree of association or predictability with other relevant measures.

During the coming months, staff will investigate how best to disaggregate its model by gender and ethnicity. Subsequent research efforts will focus on examining other possible predictors of university student success, such as the SAT, ACT, completion rates of *a–g* college preparatory curriculum, and college- and university-going rates. Once the current research phase is completed, staff will also identify measures that assess preparation for success at community colleges.

DISPLAY 2 Weighted Proficiency Measures Derived from a Principal Component Procedure

Proficiency Measure	Weight
Algebra I	.356
Geometry	.362
Integrated Math 2	.224
Algebra II	.340
World History	.372
Biology/Life Sciences	.371
Chemistry	.344
Earth Science	.299
Physics	.302

DISPLAY 3 Correlation of the Commission's Student Preparation Measure with API and University Persistence Rates

Measure		Student Preparation Measure	API
CSU	One-Year Persistence	.35	.31
	Two-Year Persistence	.35	.31
UC	One-Year Persistence	.24	.24
	Two-Year Persistence	.22	.21
API		.88	*

Technical Issues

The method used in this research endeavor is called principal component analysis. It is often used when the research focus is to summarize data and to detect linear relationships. Within the model, measures are referred to as explanatory factors.

The principal component procedure derived differential weights for various student proficiency scores of the California Standardized Test program. The weights are based on information evident in a correlation matrix, and are derived in a manner that optimizes the amount of explained variation.

The concept of variation is perhaps the most important term in the realm of educational research. Not all students or public schools exhibit a high level of proficiency with respect to a range of outcomes. Variation is a measure of the degree of difference in outcomes. A key goal of educational research is to identify a set of independent factors that explain this variation. Once these factors are better understood, appropriate improvement programs can be conceptualized and put into practice.

The principal component procedure standardized the proficiency scores into a z-score form. In this form, all proficiency scores have a common standard deviation of 1 unit, and the total performance variation is equal to the number of proficiency measures. Performance weights were determined in a manner such that the resulting linear equation would explain as much of the performance variation as possible.

Stated symbolically:

Y (Student Preparation) = .35 (Algebra I) + .36 (Geometry) + .22 (Integrated Math 2) + .34 (Algebra II) + .37 (World History) + .37 (Biology/Life Sciences) + .34 (Chemistry) + .29 (Earth Science) +.30 (Physics)

Statistically, the proficiency scores reflect the proportion of test-takers of a content domain whose scores were determined to be proficient. A numerical proficiency scale was derived by multiplying a content-specific proportion by 100. The transformation of a proportion to a numerical scale was necessary to undertake the statistical analysis.

To derive an overall student preparation score for any public high school, one would substitute the corresponding proficiency results into the equation on the preceding page.

For more information on the principal component statistical procedure, please consult *Reading and Understanding Multivariate Statistics* (Grimm & Yarnold, 1995).